

**Vergleichen Sie den SIFT- und SURF-Algorithmus hinsichtlich
Geschwindigkeit/ Komplexität und Genauigkeit/ Robustheit**

Überlegen Sie sich je eine praktische Anwendung, bei der Sie den einen Algorithmus dem anderen bevorzugen würden. Begründen Sie Ihre Wahl dieser beiden Anwendungen anhand der zuvor beschriebenen Aspekte

MARC SIMON ADAM KAMINSKI

Matr. Nr.: 03667438

TU München

18. Mai 2015

Zusammenfassung

Dieses Essay vergleicht den Scale-Invariant Feature Transform Algorithmus (SIFT) und den Speeded Up Robust Features Algorithmus (SURF), zwei Algorithmen zur Merkmalspunktextraktion, hinsichtlich Geschwindigkeit, Komplexität, Genauigkeit und Robustheit. Dazu werden wichtige Phasen der beiden Algorithmen besonders betrachtet und unter Zuhilfenahme einiger Veröffentlichungen, die sich mit dem Vergleich der beiden Algorithmen beschäftigen, ausgewertet. Mit den Ergebnissen dieses Vergleiches werden anschließend Empfehlungen für den Einsatz je eines der beiden Algorithmen in zwei verschiedenen Anwendungsgebieten ausgesprochen.

Stichpunkte: Computer Vision Essay, SIFT vs. SURF

Einleitung

Viele Anwendungen der Bildverarbeitung, wie Objekterkennung, Motion Tracking und 3D Szenenrekonstruktion, benötigen eine Erkennung markanter, lokaler Regionen im Bild, sogenannter Features oder Merkmale. Diese werden durch Detektoren aus dem Bild extrahiert, welche in verschiedenen Bildern der gleichen Szene möglichst die gleichen Features entdecken sollten. Der Detektor besitzt dann eine hohe Wiederholbarkeit (Repeatability). Im nächsten Schritt wird den gefundenen Features ein Vektor, welcher aus den Eigenschaften der Umgebung des Merkmalspunktes berechnet wird, zugewiesen. Dieser Vektor wird als Deskriptor bezeichnet und ist notwendig, um die gefundenen Features für eine Zuordnung, dem Matching, zwischen verschiedenen Bildern der gleichen Szene verwenden zu können. Deskriptoren sollten, je nach Anwendungszweck, gegenüber einer Vielzahl von Bildveränderungen, beispielsweise Rauschen, Helligkeits- und Kontraständerungen, sowie Rotations- und Verschiebeoperationen, invariant sein.

In diesem Essay werden zwei Verfahren zur Merkmalspunktextraktion, Lowes Scale-Invariant Feature Transform Algorithmus (SIFT) [1] und der Speeded Up Robust Features (SURF) Algorithmus nach Bay et al. [2], hinsichtlich ihrer Geschwindigkeit, Komplexität, Genauigkeit und Robustheit verglichen. Es wird insbesondere auf die höhere Stabilität des SIFT und die geringere Laufzeit des SURF Algorithmus eingegangen und anhand eines Beispiels aus der Zellforschung sowie aus der Überwachungstechnik verdeutlicht, welcher Algorithmus jeweils vorzuziehen wäre.

Hauptteil

SIFT und SURF Algorithmus besitzen grundsätzlich die gleiche Einteilung in Detektor, Deskriptor und Matching, verfolgen in diesen Phasen allerdings unterschiedliche Ansätze. Um skalierungsinvariante Merkmalspunkte zu erhalten, werden lokale Maxima im Skalenraum (Scalespace) berechnet. Dazu werden im SIFT Detektor verschiedene Skalierungen des Ausgangsbildes erstellt und mit gleichdimensionierten Gaußfiltern die Differenz der Mittelwerte bestimmt (Difference-of-Gaussian-Verfahren). Im Gegensatz dazu werden im SURF Detektor verschieden skalierte Gaußfilter mit dem Integralbild des Ausgangsbildes gefaltet. Die Verwendung der Integralbilder führt laut Bay et al. [2] zu einer erheblich verringerten Anzahl an FLOPS, die für die Extraktion der Merkmalspunkte nötig sind. Sie stellt allerdings auch eine Abschätzung dar. Beide Ansätze führen durch Maximumbildung schließlich zu den lokalen Maxima, welche skalierungsunabhängige Merkmalspunkte bilden. SIFT findet laut [3] prinzipiell mehr Merkmalspunkte als SURF.

Um rotations-, rausch-, helligkeits-, blickwinkel- und teilweise abbildungsinvariante Deskriptoren zu erhalten, werden nun verschiedene Filteroperationen ausgeführt. Die im SURF benutzten Operationen sind laut Bay et al.[2] dank geeigneter Approximationen schneller berechenbar als und mindestens so stabil wie für den SIFT Algorithmus. Im von Lowe [1] vorgestellten original SIFT Algorithmus führt dies zu einem 128 dimensional, im SURF Algorithmus zu einem 64 dimensional Deskriptor. Durch die geringere Dimension des SURF Deskriptors ist der Vergleich eines Merkmals mit einer Datenbank von Bildern, das sogenannte Matching, schneller durchführbar [2].

In verschiedenen Veröffentlichungen, beispielsweise [3] und [4], wurde experimentell die von Bay et al. [2] behauptete geringere Laufzeit des SURF Algorithmus bestätigt. Khan et al. [5] und Wu et al. [6] verglichen die Stabilität der beiden Algorithmen unter Anderem mit Hilfe des Quotienten aus der Anzahl richtig zugeordneter Merkmalspunkte und Anzahl der insgesamt gefundenen Merkmalspunkte, der Matchinggenauigkeit. Sie stellten dabei übereinstimmend fest, dass

der SURF Algorithmus in Hinsicht auf Skalierungsinvarianz eine schlechtere Matchinggenauigkeit aufweist als der SIFT Algorithmus. Außerdem weisen sie auf eine schlechtere Matchinggenauigkeit des SURF Algorithmus bei der Erkennung von Merkmalen in mit Unschärfe behafteten Bildern hin. Beim Vergleich der absoluten Anzahl richtig zugeordneter Merkmalspunkte durch Bauer et al. [3] wurde hingegen geschlussfolgert, dass dieser Wert beim SURF Algorithmus hoch genug für die meisten Anwendungen sei. Dabei wurde insbesondere die höhere Anzahl richtig gefundener Zuordnungen pro Zeiteinheit hervorgehoben. Demnach ist der SURF Algorithmus auf Grund seines sehr guten Kompromisses zwischen Anzahl richtig gefundener Merkmalspunkte und Laufzeit für viele Anwendungszwecke besser geeignet als SIFT.

Aus dem Vergleich folgend lassen sich die Anwendungsgebiete des SIFT und SURF Algorithmus leicht ableiten: Wird eine sehr hohe Stabilität und eine beträchtliche Anzahl an Merkmalspunkten, aber keine Echtzeitverarbeitung benötigt, ist der SIFT Algorithmus das Mittel der Wahl. Für alle Anwendungen in Echtzeitsystemen ist dazu zu raten, den SURF Algorithmus dem SIFT Algorithmus vorzuziehen.

Eine hohe Stabilität und eine große Anzahl an Merkmalspunkten ist insbesondere bei Detailaufnahmen, die beispielsweise bei der konfokalen Laser-Scanning-Mikroskopie (KLSM) von Zellen entstehen, notwendig. Mit KLSM können nicht-invasive, schnittweise 2D Aufnahmen von 3D Objekten erstellt werden. Dabei können laut Stanciu et al. [7] verschiedene Zoomlevel, Laserstrahlleistungen und andere Aufnahmeparameter gesetzt werden. Mit Hilfe von Algorithmen zur Merkmalspunktbestimmung können Merkmalspunkte in den Slices gefunden werden, um unterschiedliche Ziele zu erreichen. Zum Beispiel kann mit ihnen ein großes Bild aus vielen kleinen erstellt werden („Stitching“) oder sie werden dazu benutzt die 2D Slices aneinander auszurichten („3D scene construction“). Die Nachverarbeitung ist nicht zeitkritisch, allerdings ist eine hohe Invarianz gegenüber Bildtransformationen erforderlich um die Ausgangssituation für weitere Bildverarbeitungen zu verbessern. Ein typisches Beispiel für die Nachverarbeitung ist die 3D Zellpfadfindung. Die Verwendung des SIFT Algorithmus ist demnach für

dieses Anwendungsgebiet sinnvoll.

Ein weiteres Anwendungsgebiet von Algorithmen zur Merkmalspunktbestimmung ist die Echtzeitverarbeitung von Videoaufnahmen aus Videoüberwachungsanlagen. Auf Grund der hohen Anzahl der Bilder und um eine schnelle Reaktion von Sicherheitskräften zu ermöglichen, ist eine schnelle Verarbeitungsgeschwindigkeit unabdingbar. Ziel ist beispielsweise, wie bei Hamdoun et al. [8], die Re-Identifikation von Personen, zu denen Merkmalspunkte in einer Datenbank existieren. Die Aufnahmen werden stetig ausgewertet und gefundene Features mit der Datenbank abgeglichen. So kann die Bewegung von Personen innerhalb einer von einer Videoüberwachungsanlage überwachten Umgebung automatisch nachverfolgt werden. Um möglichst zeitnah auf identifizierte, beziehungsweise nicht identifizierte Personen in ungewöhnlichen Orten reagieren zu können, ist hier die Verwendung eines schnell arbeitenden Algorithmus wie SURF zweckhaft.

Fazit

Es wurden der SIFT und der SURF Algorithmus vorgestellt und in Hinsicht auf Komplexität, Geschwindigkeit und Stabilität verglichen. Der SIFT Algorithmus findet mehr Merkmalspunkte, ist in Hinsicht auf Skalierungs- und Unschärfelinvarianz stabiler und besitzt eine höhere Matchinggenauigkeit als der SURF Algorithmus. Dies macht ihn besonders für Bildnachverarbeitungen interessant, bei denen Wert auf eine hohe Genauigkeit gelegt wird. Der SURF Algorithmus besitzt ebenfalls eine sehr hohe Stabilität und gleichzeitig eine deutlich höhere Berechnungsgeschwindigkeit als SIFT. In Echtzeitsystemen würde ich prinzipiell den SURF-Algorithmus bevorzugen. In allen anderen Anwendungsfällen sollte eine spezielle Anwendungsprüfung beider Systeme erfolgen. Dabei sollte evaluiert werden, ob die höhere Laufzeit des SIFT Algorithmus für seine gering höhere Stabilität erkaufte werden sollte, oder auf den guten Kompromiss, den der SURF Algorithmus bietet, zurückgegriffen werden kann.

Literatur

1. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.
2. H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (surf)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.
3. J. Bauer, N. Sunderhauf, and P. Protzel, "Comparing several implementations of two recently published feature detectors," in *Proc. of the International Conference on Intelligent and Autonomous Systems*, vol. 6, 2007.
4. L. Juan and O. Gwun, "A comparison of sift, pca-sift and surf," *International Journal of Image Processing (IJIP)*, vol. 3, no. 4, pp. 143–152, 2009.
5. N. Y. Khan, B. McCane, and G. Wyvill, "Sift and surf performance evaluation against various image deformations on benchmark dataset," in *Digital Image Computing Techniques and Applications (DICTA), 2011 International Conference on*, pp. 501–506, IEEE, 2011.
6. J. Wu, Z. Cui, V. S. Sheng, P. Zhao, D. Su, and S. Gong, "A comparative study of sift and its variants," *Measurement Science Review*, vol. 13, no. 3, pp. 122–131, 2013.
7. S. G. Stanciu, R. Hristu, R. Boriga, and G. A. Stanciu, "On the suitability of sift technique to deal with image modifications specific to confocal scanning laser microscopy," *Microscopy and Microanalysis*, vol. 16, no. 05, pp. 515–530, 2010.
8. O. Hamdoun, F. Moutarde, B. Stanciulescu, and B. Steux, "Person re-identification in multi-camera system by signature based on interest point descriptors collected on short video sequences," in *Distributed Smart Cameras, 2008. ICDSC 2008. Second ACM/IEEE International Conference on*, pp. 1–6, Sept 2008.